International Journal of Statistical Distributions and Applications
2023; 9(4): 90-100
http://www.sciencepublishinggroup.com/j/ijsda
doi: 10.11648/j.ijsd.20230904.11
ISSN: 2472-3487 (Print); ISSN: 2472-3509 (Online)

Synthesis of C4 Olefin by Ethanol Coupling Based on Multivariate Statistical Analysis

Hua Xu

Department of Mathematics, Nanjing Normal University Taizhou College, Taizhou, China

Email address:
xuhua1590526@163.com

To cite this article:
Hua Xu. (2023). Synthesis of C4 Olefin by Ethanol Coupling Based on Multivariate Statistical Analysis. *International Journal of Statistical Distributions and Applications*, *9*(4), 90-100. https://doi.org/10.11648/j.ijsd.20230904.11

Received: November 12, 2023; **Accepted**: November 28, 2023; **Published**: December 11, 2023

Abstract: Ethanol coupling to prepare C4 olefin has excellent environmental benefits, economic value and broad application prospects, and has attracted extensive attention from scholars in many research fields at home and abroad. In this paper, based on the collected experimental data generated in the process of ethanol coupling to prepare C4 olefin, statistical analysis method is used to systematically analyze the data information. Firstly, regression models and exponential models of ethanol conversion, selectivity and temperature of C4 olefin were constructed, and the simulation accuracy of the two models under different catalyst combinations was compared by means of goodness of fit, significance probability and residual variance. The results showed that under 16 catalyst combinations, the ethanol conversion rate increased exponentially with the increase of temperature, and under the other 5 catalyst combinations, the ethanol conversion rate increased linearly with the increase of temperature. The selectivity of C4 olefins increased exponentially with the increase of temperature under 11 catalyst combinations, and linearly under the remaining 10 catalyst combinations. Then, based on the statistical data under temperature determination, the statistical distribution of seven different products over time was analyzed, and a multiple regression model was constructed to quantify the relationship between ethanol conversion, C4 olefin selectivity and other five products, and statistical tests were carried out. Finally, a paired sample test model was constructed to quantitatively analyze the effects of catalyst combination and temperature on ethanol conversion and selectivity of C4 olefin.

Keywords: Regression Model, Exponential Model, Accuracy Test, Statistical Test

1. Introduction

Since the reform and opening up, China has mainly relied on mineral resources to promote the development of the chemical industry, while achieving economic benefits, the excessive use of fossil energy has brought serious environmental and resource shortage problems, in the context of the realization of dual carbon goals, the development of biomass emerging energy has become the focus of today's society. As a kind of chemical raw material, C4 olefin not only requires high technology, but also consumes a lot of fossil energy under the traditional preparation method. Although China is rich in resources, but the total population is too large, the per capita energy consumption compared with the developed countries there is still a big gap, facing the most important problem is low energy efficiency, energy structure imbalance. In addition, the burning of fossil energy will produce greenhouse gases to pollute the environment, which violates the purpose of China's green and low-carbon development and is not conducive to sustainable economic development. With the innovation and progress of science and technology, the process conditions of ethanol catalytic coupling to produce C4 olefin were explored and studied. The use of ethanol as a new bioenergy can help alleviate the problem of resource shortage, reduce the consumption of traditional fossil fuels, and help maintain ecological harmony. C4 olefin is an important raw material for chemical industry products and medicine, and efficient access to more C4 olefin can promote the development of chemical industry, formulate more drugs, meet the needs of society, and maintain economic development and stability. As a kind of clean energy and a wide range of sources, ethanol is conducive to slowing down the trend of global warming. Ethanol coupling preparation of C4 olefin has excellent environmental benefits and broad application prospects, and has been widely concerned by scholars from industry, environment, economy and other research fields.

The preparation of C4 olefin from ethanol usually requires dehydrogenation and coupling. In the preparation process, different catalyst combinations and temperature control during the preparation process will have different degrees of influence on the selectivity and yield of C4 olefin. Therefore, in practice, it is of great practical significance to carefully design different types of catalyst combinations and study the catalytic coupling of ethanol to produce C4 olefin under different environmental conditions. In recent years, domestic scholars have carried out a series of studies on the production of C4 olefin by ethanol catalytic coupling. For example, Fang Yangyang et al. [1] screened the factors affecting the production of C4 olefin by ethanol catalytic coupling based on the traditional correlation measurement method, and discussed the effect of catalyst combination on the yield of C4 olefin. Wu Wenjun et al. [2] analyzed the experimental data by constructing RBF neural network and quantitatively evaluated the influence of different catalyst combinations on C4 olefin selectivity by weighting model neurons. Li Sanjie et al. [3] used the research method of grey system theory to divide the experimental data into reference sequence and contrast sequence, calculate the grey correlation degree between them, and then seek the optimization scheme of C4 olefin conversion rate. Wang Runmo et al. [4], based on principal component analysis and partial least squares regression analysis, established a multivariate nonlinear optimization model considering various constraints to determine the best combination and collocation. Wang Liying et al. [5] analyzed the factors affecting the conversion rate of C4 olefin from various aspects based on MATLAB and computer related knowledge. Zhang Yuan et al. [6] firstly optimized the process conditions by using neural networks and multivariate nonlinear fitting, and proposed models of ethanol conversion rate and C4 olefin selectivity and temperature based on Arrhenius experiment. The relevant regression equation is obtained by fitting, and the optimal model is established to obtain the theoretical maximum charge rate. Jiang Zihao et al. [7] used the analysis of variance to analyze the important influence factors, then analyzed the importance ranking of the influence on the combination, and finally carried out multivariate nonlinear regression through the L-M algorithm to obtain the optimal yield. Bo Xiaohan et al. [8] analyzed the correlation using the least square fitting, discussed the influence of different variables using Spearman correlation coefficient, and then established a multiple quadratic regression model to compare the differences in charging methods, and finally designed a reasonable scheme. Li Shaowei et al. [9] proposed an exponential regression equation model based on Arrhenius equation, then analyzed the important factors affecting the transformation by using grey correlation, and finally established an optimization model by using Gaussian process regression. Li Ming et al. [10] discussed the reaction degree of preparation of C4 olefin by the combination of various catalysts through regression analysis, so as to select the best catalyst combination. Zhang et al. [11] studied the ethanol conversion rate and C4 olefin selectivity, respectively, and used the ethanol conversion rate as an example to illustrate. Shi [12] used SMO algorithm to study that the yield of C4 olefins is the highest when the temperature is 350℃. Zhang [13] established a multivariate linear regression model to fit the ethanol conversion rate and C4 olefin selectivity. Pang et al. [14] used partial correlation analysis and linear stepwise regression to Study on preparation of olefins. Chen [15] explored the relationship between selectivity of C4 olefin and catalyst combination and temperature, and how to optimize catalyst combination and temperature to make the yield of C4 olefin as high as possible.

In summary, we can see that more systematic studies have been carried out at home and abroad on the catalytic coupling of ethanol to produce C4 olefin. The research focuses on the modeling of the relationship between ethanol conversion rate, C4 olefin selectivity, catalyst combination and temperature. The research methods adopted include regression analysis, neural network, analysis of variance, grey correlation analysis, etc. And a series of research results have been obtained. However, we found that there are still many shortcomings in the existing studies. For example, the existing literature is often based on a specific method (such as regression model) to describe the relationship between ethanol conversion rate and C4 olefin selectivity and temperature, but in fact, there is still a lot of room for improvement by using only one model under different catalyst combinations. Second, the existing studies still lack the quantitative characterization of the relationship between ethanol conversion rate and C4 olefin selectivity and other products, but in fact, there are complex internal relationships between ethanol conversion rate and C4 olefin selectivity and other products in the process of ethanol catalytic coupling to produce C4 olefin. Based on the above analysis, it can be seen that there are still many problems in the preparation of C4 olefin by ethanol catalytic coupling that need further research.

Based on the experimental data, this paper focuses on the following research contents: (1) The quantitative relationship between the ethanol conversion rate and the selectivity of C4 olefin under different catalyst combinations and temperature is studied. Based on the experimental data, the experimental data of ethanol conversion rate and C4 olefin selectivity under each catalyst combination is calculated and visualized with MATLAB. On this basis, regression models and exponential models of ethanol conversion, selectivity and temperature of C4 olefin were constructed. The simulation accuracy of the two types of models under different catalyst combinations was compared by means of goodness-fit, significance probability, residual variance, etc., and reference suggestions for model selection under 21 catalyst combinations were given. (2) The quantitative relationship between ethanol conversion rate, C4 olefin selectivity and other products was studied. Based on the statistical data under temperature determination, the statistical distribution law of seven different products over time was analyzed, and a multiple regression model was constructed to quantitatively analyze the relationship between ethanol conversion rate, C4 olefin selectivity and other products and statistical test was conducted.

2. Modeling and Result Analysis

2.1. Data Sources

The data used in this paper comes from the experimental data provided by the B question of 2021 Higher Education Society Cup National College Students Mathematical Contest in Modeling (http://www.mcm.edu.cn/html_cn/node/90d223833c1eb50f899aa096a66c6896.html).

The experiment provides quantitative data on the relationship between ethanol conversion and C4 olefin selectivity and temperature under 21 different catalyst combinations, among which 21 different catalyst combinations are shown in the table below:

Table 1. *Catalyst combination number and catalyst combination.*

Number	A1	A2	A3
Catalyst combination	200mg 1wt%Co/SiO2 - 200mg HAP-ethanol concentration 1.68ml/min	200mg 2wt%Co/SiO2- 200mg HAP-ethanol concentration 1.68ml/min	200mg 1wt%Co/SiO2- 200mg HAP-ethanol concentration 0.9ml/min
Number	A4	A5	A6
Catalyst combination	200mg 0.5wt%Co/SiO2- 200mg HAP-ethanol concentration 1.68ml/min	200mg 2wt%Co/SiO2- 200mg HAP-ethanol concentration 0.3ml/min	200mg 5wt%Co/SiO2- 200mg HAP-ethanol concentration 1.68ml/min
Number	A7	A8	A9
Catalyst combination	50mg 1wt%Co/SiO2- 50mg HAP- ethanol concentration 0.3ml/min	50mg 1wt%Co/SiO2- 50mg HAP- ethanol concentration 0.9ml/min	50mg 1wt%Co/SiO2- 50mg HAP-ethanol concentration 2.1ml/min
Number	A10	A11	A12
Catalyst combination	50mg 5wt%Co/SiO2- 50mg HAP- ethanol concentration 2.1ml/min	50mg 1wt%Co/SiO2+ 90mg- ethanol concentration 1.68ml/min, No HAP	50mg 1wt%Co/SiO2- 50mg HAP-ethanol concentration 1.68ml/min
Number	A13	A14	B1
Catalyst combination	67mg 1wt%Co/SiO2- 33mg HAP- ethanol concentration 1.68ml/min	33mg 1wt%Co/SiO2- 67mg HAP-ethanol concentration 1.68ml/min	50mg 1wt%Co/SiO2- 50mg HAP-ethanol concentration 1.68ml/min
Number	B2	B3	B4
Catalyst combination	100mg 1wt%Co/SiO2- 100mg HAP-ethanol concentration 1.68ml/min	10mg 1wt%Co/SiO2- 10mg HAP-ethanol concentration 1.68ml/min	25mg 1wt%Co/SiO2- 25mg HAP-ethanol concentration 1.68ml/min
Number	B5	B6	B6
Catalyst combination	50mg 1wt%Co/SiO2- 50mg HAP- ethanol concentration 2.1ml/min	75mg 1wt%Co/SiO2- 75mg HAP-ethanol concentration 1.68ml/min	100mg 1wt%Co/SiO2- 100mg HAP-ethanol concentration 0.9ml/min

2.2. Modeling and Analysis of the Relationship Between Ethanol Conversion and Temperature

Based on the experimental data, we can see that the ethanol conversion rate presents an increasing trend with the increase of temperature. Under the combination of catalysts A2, A3, A4, A6 and A7, the ethanol conversion rate presents a linear increase trend with the increase of temperature. In the case of A1, A5, A8, A9, A10, A11, A12, A13, A14, B1, B2, B3, B4, B5, B6 and B7 catalyst combination, the ethanol conversion rate increased exponentially with the increase of temperature. For this reason, we constructed a unitary linear regression model under each catalyst combination, respectively, with ethanol conversion as y and temperature as x, as follows:

$$y = b_0 + b_1 x + \varepsilon, \quad (1)$$

Where b_0 and b_1 are regression coefficients and ε is the error term.

The exponential growth model is constructed as follows:

$$y = \beta_0 e^{\beta_1 x}, \quad (2)$$

To solve the exponential growth model, we convert it to linear form:

$$lny = ln\beta_0 + \beta_1 x, \quad (3)$$

To do this, we can estimate the model parameters β_0 and β_1 with the help of linear least squares.

The parameter estimation results of model (1) and model (2) and the accuracy test results of the model under 21 different catalyst combinations were obtained through MATLAB programming, as shown in Table 2.

Table 2. *Estimation results of the relationship between ethanol conversion and temperature under different catalyst combinations.*

Catalyst combination	Mathematical model	Coefficient	Estimatedvalue	Confidence interval	Accuracy test
A1	Linear regression	b_0	-84.0740	[-133.9694, -34.1786]	R^2=0.9321, p=0.0077
		b_1	0.3332	[0.1680, 0.4983]	
	Exponential model	β_0	0.0025	[0.0001, 0.0640]	R^2=0.9576, p=0.0038
		β_1	0.0279	[0.0171, 0.0387]	
A2	Linear regression	b_0	-161.8920	[-198.8224 -124.9616]	R^2=0.9900, p=0.0004
		b_1	0.6630	[0.5407, 0.7852]	
	Exponential model	β_0	0.0098	[0.0001, 1.6697]	R^2=0.8895, p=0.0161
		β_1	0.0263	[0.0093, 0.0433]	

Catalyst combination	Mathematical model	Coefficient	Estimatedvalue	Confidence interval	Accuracy test
A3	Linear regression	b_0	-95.8275	[-127.5650, -64.0901]	R^2=0.9643, p=0.0001
		b_1	0.4194	[0.3266, 0.5122]	
	Exponential model	β_0	1.0138	[0.2495, 4.1201]	R^2=0.8983, p=0.0012
		β_1	0.0106	[0.0065, 0.0147]	
A4	Linear regression	b_0	-144.5400	[-162.8186, -126.2614]	R^2=0.9950, p=0.0000
		b_1	0.5816	[0.5246, 0.6386]	
	Exponential model	β_0	0.0496	[0.0017, 1.4809]	R^2=0.8707, p=0.0066
		β_1	0.0198	[0.0092, 0.0304]	
A5	Linear regression	b_0	-97.5929	[-166.7687, -28.4170]	R^2=0.8730, p=0.0063
		b_1	0.4077	[0.1919, 0.6236]	
	Exponential model	β_0	0.5947	[0.1639, 2.1577]	R^2=0.9442, p=0.0012
		β_1	0.0119	[0.0079, 0.0159]	
A6	Linear regression	b_0	-119.7310	[-173.6412, -65.8209]	R^2=0.9675, p=0.0025
		b_1	0.5012	[0.3325, 0.6699]	
	Exponential model	β_0	0.3972	[0.0661, 2.3864]	R^2=0.9521, p=0.0045
		β_1	0.0136	[0.0080, 0.0192]	
A7	Linear regression	b_0	-74.1819	[-82.0209, -66.3429]	R^2=1.0000, p=0.0000
		b_1	0.3773	[0.3527, 0.4018]	
	Exponential model	β_0	2.5409	[0.8218, 7.8565]	R^2=0.9539, p=0.0043
		β_1	0.0087	[0.0052, 0.0123]	
A8	Linear regression	b_0	-83.6302	[-127.0364, -40.2240]	R^2=0.9547, p=0.0042
		b_1	0.3392	[0.2034, 0.4750]	
	Exponential model	β_0	0.1468	[0.0747, 0.2883]	R^2=0.9942, p=0.0002
		β_1	0.0150	[0.0129, 0.0171]	
A9	Linear regression	b_0	-65.6241	[-127.6373, -3.6110]	R^2=0.8475, p=0.0265
		b_1	0.2490	[0.0549, 0.4430]	
	Exponential model	β_0	0.0126	[0.0056, 0.0281]	R^2=0.9953, p=0.0001
		β_1	0.0200	[0.0175, 0.0226]	
A10	Linear regression	b_0	-49.6879	[-94.6255, -4.7504]	R^2=0.8519, p=0.0254
		b_1	0.1835	[0.0429, 0.3241]	
	Exponential model	β_0	0.0002	[0.0000, 0.0016]	R^2=0.9876, p=0.0006
		β_1	0.0298	[0.0237, 0.0359]	
A11	Linear regression	b_0	-56.5741	[-114.3165, 1.1682]	R^2=0.8158, p=0.0356
		b_1	0.2070	[0.0263, 0.3877]	
	Exponential model	β_0	0.00004	[0.0000, 0.0002]	R^2=0.9922, p=0.0003
		β_1	0.0341	[0.0286, 0.0397]	
A12	Linear regression	b_0	-74.7957	[-121.6123, -27.9791]	R^2=0.9278, p=0.0084
		b_1	0.2858	[0.1393, 0.4323]	
	Exponential model	β_0	0.0066	[0.0009, 0.0496]	R^2=0.9771, p=0.0015
		β_1	0.0225	[0.0162, 0.0288]	
A13	Linear regression	b_0	-67.3328	[-122.9906, -11.6749]	R^2=0.8772, p=0.0190
		b_1	0.2533	[0.0791, 0.4275]	
	Exponential model	β_0	0.0040	[0.0023, 0.0069]	R^2=1.0000, p=0.0000
		β_1	0.0232	[0.0214, 0.0249]	
A14	Linear regression	b_0	-86.6440	[-141.1240, -32.1640]	R^2=0.9291, p=0.0082
		b_1	0.3358	[0.1653, 0.5062]	
	Exponential model	β_0	0.0212	[0.0043, 0.1042]	R^2=0.9817, p=0.0011
		β_1	0.0199	[0.0149, 0.0249]	
B1	Linear regression	b_0	-73.1897	[-119.7940, -26.5853]	R^2=0.9254, p=0.0088
		b_1	0.2796	[0.1338, 0.4254]	
	Exponential model	β_0	0.0066	[0.0010, 0.0446]	R^2=0.9793, p=0.0013
		β_1	0.0224	[0.0164, 0.0284]	
B2	Linear regression	b_0	-70.8595	[-134.4954, -7.2235]	R^2=0.8633, p=0.0224
		b_1	0.2724	[0.0732, 0.4715]	
	Exponential model	β_0	0.0265	[0.0145, 0.0485]	R^2=0.9969, p=0.0001
		β_1	0.0185	[0.0166, 0.0203]	
B3	Linear regression	b_0	-36.1843	[-66.1375, -6.2311]	R^2=0.7920, p=0.0175
		b_1	0.1314	[0.0379, 0.2248]	
	Exponential model	β_0	0.0003	[0.0001, 0.0013]	R^2=0.9882, p=0.0001
		β_1	0.0277	[0.0235, 0.0319]	
B4	Linear regression	b_0	-56.9000	[-101.8311, -11.9689]	R^2=0.8092, p=0.0146
		b_1	0.2080	[0.0678, 0.3482]	
	Exponential model	β_0	0.0005	[0.0001, 0.0025]	R^2=0.9856, p=0.0001
		β_1	0.0279	[0.0232, 0.0326]	
B5	Linear regression	b_0	-72.4000	[-126.5177, -18.2823]	R^2=0.8333, p=0.0111
		b_1	0.2720	[0.1031, 0.4409]	
	Exponential model	β_0	0.0140	[0.0106, 0.0183]	R^2=1.0000, p=0.0000
		β_1	0.0202	[0.0193, 0.0210]	
B6	Linear regression	b_0	-99.8829	[-161.8235, -37.9422]	R^2=0.8835, p=0.0053
		b_1	0.3833	[0.1900, 0.5766]	
	Exponential	β_0	0.0294	[0.0058, 0.1490]	R^2=0.9660, p=0.0004

Catalyst combination	Mathematical model	Coefficient	Estimatedvalue	Confidence interval	Accuracy test
	model	β_1	0.0194	[0.0144 0.0245]	
B7	Linear regression	b_0	-109.3429	[-179.4934, -39.1924]	R^2=0.8763, p=0.0060
		b_1	0.4197	[0.2008, 0.6386]	
	Exponential model	β_0	0.0498	[0.0330, 0.0752]	R^2=1.0000, p=0.0000
		β_1	0.0182	[0.0169, 0.0195]	

It can be seen from the calculation results in Table 2 that under the catalyst combination A2, A3, A4, A6 and A7, the simulation result of linear model (1) is better than that of exponential model (2), indicating that under the catalyst combination A2, A3, A4, A6 and A7, the ethanol conversion rate increases linearly with the increase of temperature. In the case of catalyst combination A2, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.6630x - 161.8920$, goodness of fit $R^2 = 0.9900$, significance probability $p = 0.0004 < 0.01$; In the case of A3, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.4194x - 95.8275$, goodness of fit $R^2 = 0.9643$, $p = 0.0001 < 0.01$; Under A4 condition, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.5816x - 144.54$, $R^2 = 0.9950, p = 0.0000 < 0.01$; In the case of A6, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.5012x - 119.7310$, $R^2 = 0.9675, p = 0.0025 < 0.01$; In the case of A7, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.3773x - 74.1819$, goodness of fit $R^2 = 1.0000$, and significance probability $p = 0.0000 < 0.01$.

In the case of catalyst combination A1, A5, A8, A9, A10, A11, A12, A13, A14, B1, B2, B3, B4, B5, B6 and B7, the simulation result of exponential model (2) is better than that of linear model (1). It is indicated that under the conditions of catalyst combination A1, A5, A8, A9, A10, A11, A12, A13, A14, B1, B2, B3, B4, B5, B6 and B7, the ethanol conversion rate increases exponentially with the increase of temperature. In the case of catalyst combination A1, the quantitative relationship between ethanol conversion rate (%) and temperature (°C) is expressed as follows: $y = 0.0025e^{0.0279x}$, goodness of fit $R^2 = 0.9576$, significance probability $p = 0.0038 < 0.01$; In the case of A5, the quantification relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.5947e^{0.0119x}$, $R^2 = 0.9442$, and the significance probability $p = 0.0012 < 0.01$; In the case of A8, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.1468e^{0.0150x}$, goodness of fit $R^2 = 0.9942$, significance probability $p = 0.0002 < 0.01$; In the case of A9, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0126e^{0.0200x}$, $R^2 = 0.9953$, significance probability $p = 0.0001 < 0.01$; In the case of A10, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0002e^{0.0298x}, R^2 = 0.9876, p = 0.0006 < 0.01$; In the case of A11, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.00004e^{0.0341x}$, goodness of fit $R^2 = 0.9922$, significance probability $p = 0.0003 < 0.01$; In the case of A12, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0066e^{0.0225x}, R^2 = 0.9771, p = 0.0015 < 0.01$; In the case of A13, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0040e^{0.0232x}$, goodness of fit $R^2 = 1.0000$, significance probability $p = 0.0000 < 0.01$; In the case of A14, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0212e^{0.0199x}, R^2 = 0.9817, p = 0.0011 < 0.01$; In the case of B1, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0066e^{0.0224x}$, goodness of fit $R^2 = 0.9793$, significance probability $p = 0.0013 < 0.01$; In the case of B2, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0265e^{0.0185x}, R^2 = 0.9969, p = 0.0001 < 0.01$; In the case of B3, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0003e^{0.0277x}$, goodness of fit $R^2 = 0.9882$, significance probability $p = 0.0001 < 0.01$; In the case of B4, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0005e^{0.0279x}$, goodness of fit $R^2 = 0.9856$, significance probability $p = 0.0001 < 0.01$; In the case of B5, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0140e^{0.0202x}, R^2 = 1.0000, p = 0.0000 < 0.01$; In the case of B6, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0294e^{0.0194x}, R^2 = 0.9660, p = 0.0004 < 0.01$; In the case of B7, the quantization relationship between ethanol conversion (%) and temperature (°C) is as follows: $y = 0.0498e^{0.0182x}$, goodness of fit $R^2 = 1.0000$, and significance probability $p = 0.0000 < 0.01$.

In summary, it can be seen from the modeling analysis that under different catalyst combinations, the ethanol conversion rate presents a trend of linear or exponential growth with the change of temperature. The linear or exponential function can be used to establish the model relationship between the ethanol conversion rate and the change of temperature, and the model can achieve good simulation accuracy.

2.3. Modeling and Analysis of the Relationship Between Selectivity and Temperature of C4 Olefin

Under 19 catalyst combinations, such as A2, A4, A5-A14 and B1-B7, the selectivity of C4 olefins (%) increased with the increase of temperature (°C). Therefore, we can still choose the model (1) and model (2) constructed above to model and analyze the relationship between C4 olefin selectivity (%) and temperature (°C). In the case of catalyst combination A1 and A3, we found that the selectivity of C4 olefin first increased and then decreased with the increase of temperature (°C). Obviously, at this time, neither model (1) nor model (2) can well reflect the relationship between C4 olefin selectivity (%) and temperature (°C). Therefore, for the catalyst combination A1 and A3, we consider building the following quadratic regression model as follows:

$$y = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \varepsilon, \quad (4)$$

Where α_0, α_1 and α_2 are regression coefficients, and ε is the error term.

The results of parameter estimation and accuracy test of model (1), model (2) and model (4) under 21 different catalyst combinations were obtained by MATLAB programming. It can be seen from the calculated results that in the case of catalyst combination A1, the goodness of fit R^2 of the primary regression model is 0.7869, and the significance probability $p = 0.0448$; the goodness of fit R^2 of the exponential model is 0.7935, and the significance probability $p = 0.0426$. The accuracy of the modeling effect of the primary regression model and the exponential model needs to be further improved, so we build a quadratic regression model:

$$y = -0.0021x^2 + 1.4222x - 190.7834,$$

At this time, the goodness of fit $R^2 = 0.9160$ and the significance probability $p = 0.0084$ of the model, the fitting accuracy of the quadratic model has been significantly improved compared with the linear regression and exponential model. In the case of catalyst combination A2, the exponential model results are better than the linear model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 1.8634e^{0.0085x}$, good of fit $R^2 = 0.8517$, significance probability $p = 0.0254$. The situation in the case of A3 is similar to that in the case of A1, the simulation accuracy of the quadratic regression model is significantly better than that of the primary model and the exponential model, and the quadratic regression model of the relationship between C4 olefin selectivity (%) and temperature (°C) in the case of A3 is obtained:

$$y = -0.0009x^2 + 0.9244x - 171.0759,$$

The goodness of fit $R^2 = 0.9551$, the significance probability $p = 0.0020$; In the case of A4, the results of the exponential model are better than those of the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.4636e^{0.0112x}$, goodnessof fit $R^2 = 0.9229$, and significance probability $p = 0.0023$. In the case of A5, the results of the linear model are better than those of the exponential model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.2297x - 57.8127$, the goodness of fit $R^2 = 0.9401$, and the probability of significance $p = 0.0014$. In the case of A6, the results of the exponential model are superior to those of the linear model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1076e^{0.0141x}$, goodness of fit $R^2 = 0.9155$, and significance probability $p = 0.0107$. In the case of A7, the results of the exponential model are better than those of the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) can be obtained as follows: $y = 0.2388e^{0.0123x}$, goodness of fit $R^2 = 0.9887$, significance probability $p = 0.0005$. In the case of A8, the linear model results are better than the exponential model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.2423x - 57.2566$, goodness of fit $R^2 = 0.9832$, and significance probability $p = 0.0009$. In the case of A9, the linear model results are better than the exponential model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.2538x - 59.0950$, good of fit $R^2 = 0.9948$, and significance probability $p = 0.0002$. In the case of A10, the results of the exponential model are superior to those of the linear model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1002e^{0.0109x}$, goodness of fit $R^2 = 0.8176$, and significance probability $p = 0.0351$. In the case of A11, the linear model results are better than the exponential model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.0519x - 13.3074$, goodness of fit $R^2 = 0.9782$, and significance probability $p = 0.0014$. In the case of A12, the results of the exponential model are better than those of the linear model, and the quantitative relationship between the C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.2981e^{0.0121x}$, the goodness of fit $R^2 = 0.9962$, and the probability of significance $p = 0.0001$. In the case of A13, the results of the linear model are better than those of the exponential model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1628x - 35.8893$, goodness of fit $R^2 = 0.9768$, and significance probability $p = 0.0015$. In the case of A14, the results of the exponential model are superior to those of the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) can be obtained as follows: $y = 0.0236e^{0.0172x}$, goodness of fit $R^2 = 0.9911$, and significance probability $p = 0.0004$. In summary, it can be seen that under the catalyst combination A1, A3, A5, A8, A9, A11 and A13, the modeling accuracy of the quantitative relationship between C4 olefin selectivity and temperature based on regression model is better than that of the exponential model. Under the catalyst combination A2, A4, A6, A7, A10, A12 and A14, the modeling accuracy of the quantitative relationship between C4 olefin selectivity and temperature based on the exponential model is better than that of the regression model.

We continue to discuss the quantitative relationship between C4 olefin selectivity and temperature in the case of catalyst combination B1-B7. In the case of B1, the results of the exponential model are superior to the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) can be expressed as follows: $y = 0.2476e^{0.0130x}$, goodness of fit $R^2 = 0.9908$, significance probability $p = 0.0004$; In the case of B2, the exponential model results are better than the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.0508e^{0.0170x}$, good of fit $R^2 = 0.9811$, significance probability $p = 0.0011$. In the case of B3, the exponential model results are better than the linear model, and the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1423e^{0.0127x}$, goodness of fit $R^2 = 0.9525$, significance probability $p = 0.0009$. In the case of B4, the result accuracy of the linear model is higher than that of the exponential model, but the confidence interval of the coefficient contains zero, so the exponential model is still chosen. Thus, the quantitative relationship between C4 olefin selectivity (%) and temperature (°C) can be expressed as follows: $y = 0.1023x - 22.2081$, goodness of fit $R^2 = 0.8081$, significance probability $p = 0.0160$; In the case of B5, the results of the exponential model are better than those of the linear model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1804e^{0.0126x}$, good of fit $R^2 = 0.9874$, and significance probability $p = 0.0001$. In the case of B6, the results of the linear model are better than those of the exponential model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.1903x - 45.7573$, goodness of fit $R^2 = 0.9646$, and significance probability $p = 0.0005$. In the case of B7, the results of the linear model are better than those of the exponential model, and the quantization relationship between C4 olefin selectivity (%) and temperature (°C) is as follows: $y = 0.2337x - 56.4513$, goodness of fit $R^2 = 0.9888$, and significance probability $p = 0.0000$. In summary, it can be seen that under the catalyst combination B4, B6 and B7, the modeling accuracy of the quantitative relationship between C4 olefin selectivity and temperature based on regression model is better than that of the exponential model. In the case of catalyst combination B1, B2, B3 and B5, the modeling accuracy of the quantitative relationship between C4 olefin selectivity and temperature based on the exponential model is better than that of the regression model.

2.4. Modeling and Analysis of the Relationship Between Ethanol Conversion, C4 Olefin Selectivity and Other Products Under Temperature Determination

Based on the experimental data at 350°C, we used MATLAB programming to draw the changing trends of seven products including ethanol conversion, C4 olefin selectivity, ethylene selectivity, acetaldehyde selectivity, fatty alcohols with carbon number 4-12, methylbenzaldehyde and methylbenzyl alcohol over time. And the variation trend of C4 olefin yield over time was obtained through calculation, as shown in Figure 1.

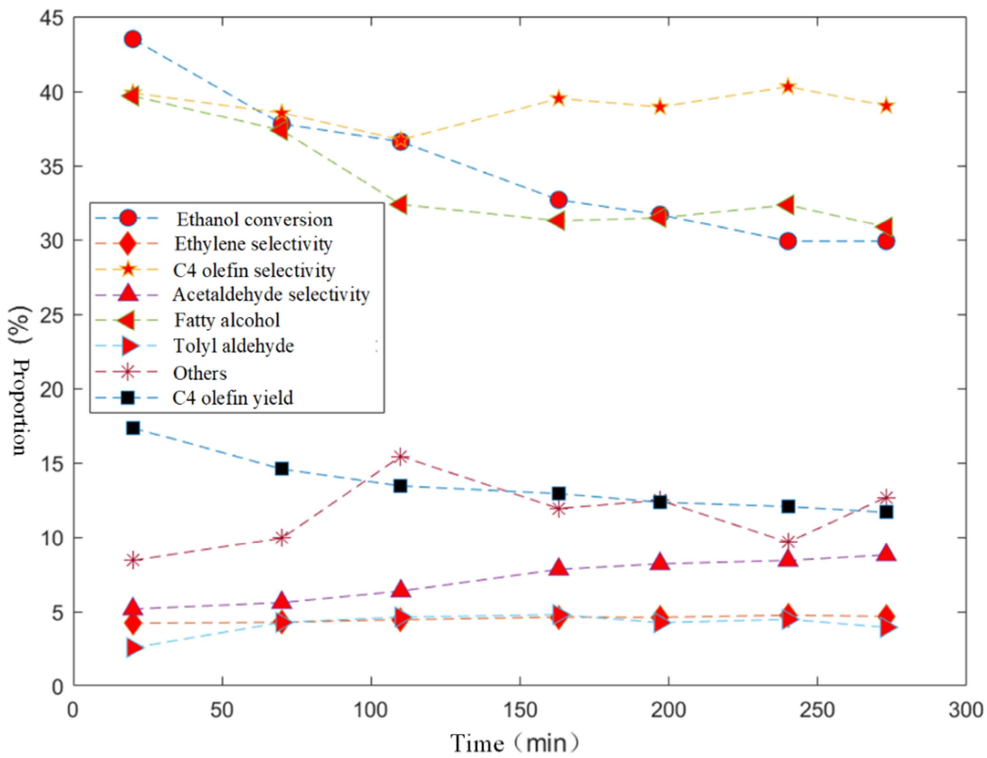

Figure 1. *Image of the change of test data over time for a given catalyst combination at 350°C.*

The calculated statistical distribution images of each indicator data are shown in Figure 2, and the specific descriptive statistical results are shown in Table 3.

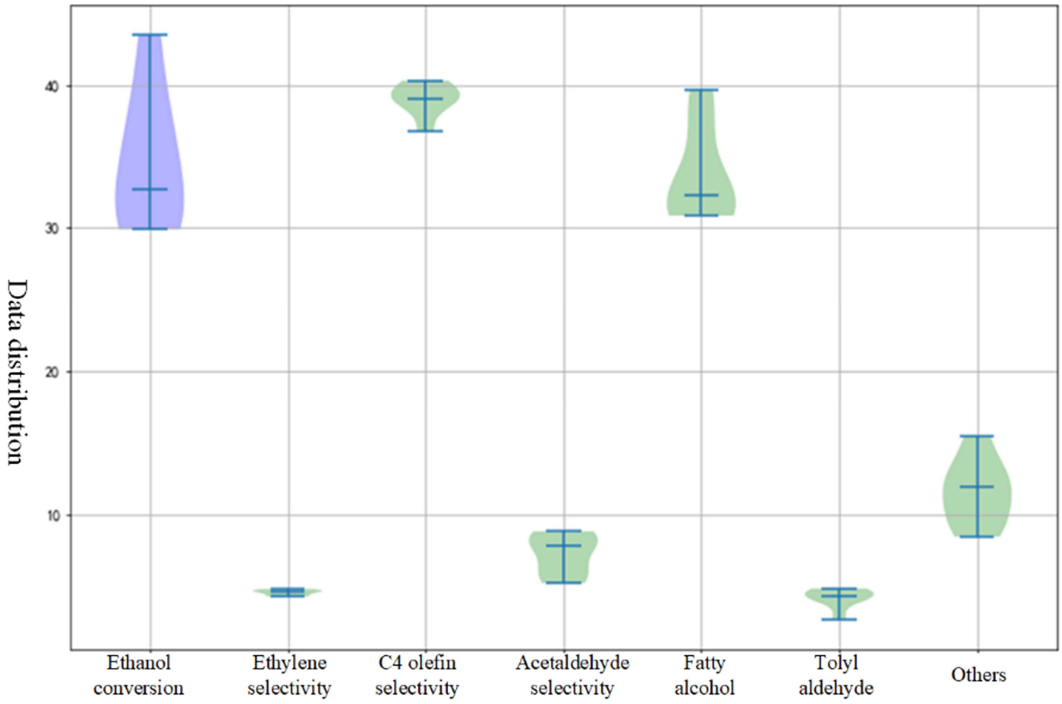

Figure 2. *Statistical image of data distribution.*

Table 3. *Descriptive statistical results of data.*

Various indexes	Minimum value	Maximum value	Average value	Standard deviation	Variance
	Statistics	Statistics	Statistics	Statistics	Statistics
Ethanol conversion	29.9	43.5	34.583	5.0080	25.080
Ethylene selectivity	4.23	4.76	4.5229	0.20435	0.042
C4 olefin selectivity	36.72	40.32	39.0029	1.17170	1.373
Acetaldehyde selectivity	5.17	8.79	7.1943	1.45777	2.125
The carbon number is 4-12 fatty alcohols	30.86	39.70	33.6357	3.45534	11.939
Methylbenzaldehyde and methylbenzyl alcohol	2.58	4.80	4.1400	0.74106	0.549
Else	8.42	15.43	11.5043	2.35653	5.553

	Measure of skewness		Kurtosis	
	Statistics	Standard error	Statistics	Standard error
Ethanol conversion	0.971	0.794	0.265	1.587
Ethylene selectivity	-0.556	0.794	-1.427	1.587
C4 olefin selectivity	-1.282	0.794	2.259	1.587
Acetaldehyde selectivity	-0.435	0.794	-1.929	1.587
The carbon number is 4-12 fatty alcohols	1.281	0.794	0.037	1.587
Methylbenzaldehyde and methylbenzyl alcohol	-1.923	0.794	4.171	1.587
Else	0.421	0.794	-0.125	1.587

According to the calculation results in Figure 1, Figure 2 and Table 3, it can be seen that the conversion rate of ethanol shows a decreasing trend over time. On the whole, the mean value of ethanol conversion is 34.59%, the standard deviation is 4.99, the maximum value is 43.5% at 20min, and the stable value is 29.9% after 240min. Ethylene selectivity showed an upward trend with time. On the whole, the mean value of ethylene selectivity was 4.52%, the standard deviation was 0.20, and the maximum value was 4.76% at 240min. The selectivity of C4 olefin decreased first and then increased with time, the mean value was 39%, the standard deviation was 1.17, the minimum value appeared at 110min was 36.72%, and the maximum appeared at 240min was 40.32. Acetaldehyde selectivity showed an increasing trend with time, the mean value was 7.19% and the standard deviation was 1.46. The mean value and standard deviation of 4-12 fatty alcohols were 33.64% and 3.46 respectively. Methylbenzaldehyde and methylbenzyl alcohol increased first and then decreased with the time, the mean value was 4.14 and the standard deviation was 0.74. The other products fluctuate over time with a mean of 11.50 and a standard deviation of 2.36. From the coefficient of variation of each index data, C4 olefin selectivity has the smallest coefficient of variation, indicating that compared with other indexes, C4 olefin selectivity has the smallest variation over time and is relatively stable. The coefficient of variation of other products is the largest, indicating that compared with other indicators, other products have the largest range of change over time and are the most unstable. As can be seen from Figure 3, when the reaction temperature reaches 350 degrees, when the reaction time enters 20min, the yield of C4 olefin reaches the maximum value, and the corresponding value is 17.38%. After reaching the peak value, the yield of C4 olefin begins to gradually decline with the passing of time. Controlling the reaction time can effectively improve the yield of C4 olefin, and then improve the efficiency of preparing C4 olefin.

In order to further explore the quantitative relationship among five products, such as ethanol conversion, C4 olefin selectivity, ethylene selectivity and acetaldehyde selectivity, the following multiple regression model was established:

$$\begin{cases} y_1 = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + \alpha_5 x_5 + \varepsilon_1 \\ y_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon_2 \end{cases} \quad (5)$$

Wherey_1 represents ethanol conversion, y_2 represents C4 olefin selectivity, $x_1, \dots, x_5$ respectively represents ethylene selectivity, acetaldehyde selectivity, carbon number 4-12 fatty alcohols, methylbenzaldehyde and other products, $\alpha_0, \dots, \alpha_5, \beta_0, \dots, \beta_5$ represents the regression coefficient, and ε_1 and ε_2 are random terms.

Based on the data of ethanol conversion and C4 olefin selectivity obtained above, the quantization results between ethanol conversion and five products were calculated by programming according to formula (5), as shown in Table 4.

Table 4. *Calculation results of the relationship between ethanol conversion and five products.*

Regression coefficient	Coefficient estimates	Coefficient confidence interval
α_0	197.9882	[134.2216, 530.1980]
α_1	-2.8170	[-38.1254, -2.4913]
α_2	-6.2367	[-12.6876, -0.2143]
α_3	-2.2973	[-6.8744, -2.2799]
α_4	-3.8717	[-7.9881, -0.2447]
α_5	-1.0860	[-3.7131, -1.0411]
$R^2 = 0.9996,\ F = 527.5409,\ p < 0.05,\ s^2 = 0.0567$		

From Table 4, the estimate of y_1 can be expressed as

$$\hat{y}_1 = 197.9882 - 2.8170x_1 - 6.2367x_2 - 2.2973x_3 - 3.8717x_4 - 1.0860x_5$$

The comparison image between the estimated value of y_1 and the actual value and the confidence interval of the model residual are shown in Figure 3.

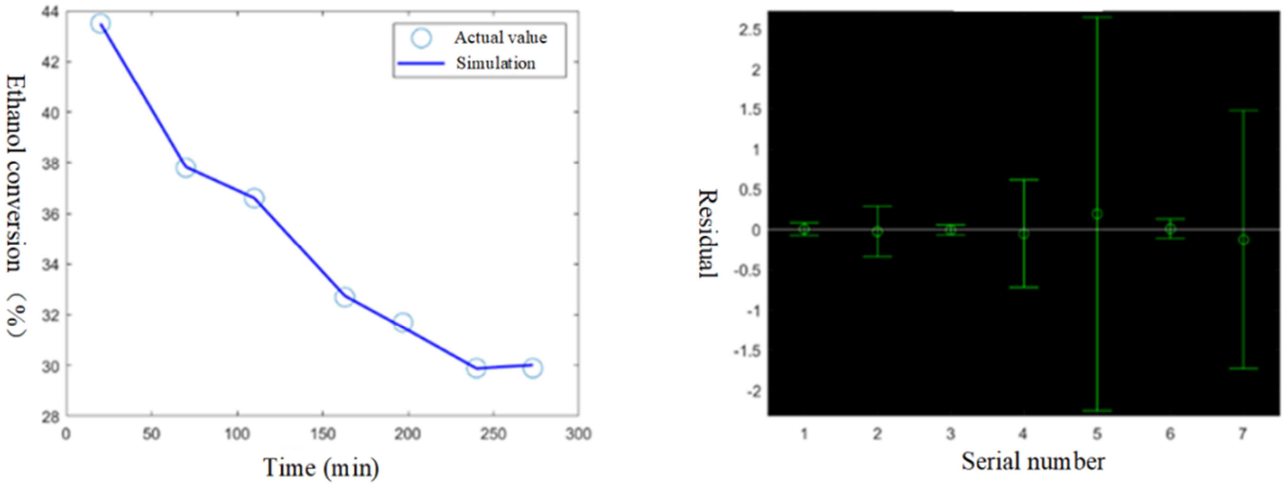

Figure 3. *Comparison image between the estimated value of y_1 and the actual value and the model residual confidence interval.*

The quantitative relationship between C4 olefin selectivity and the five products was obtained by the same method:

$$\hat{y}_2 = 100 - 0.9897x_1 - 0.9988x_2 - 0.9623x_3 - 0.9578x_4 - 0.9943x_5$$

The comparison image between the estimated value of y_2 and the actual value and the residual confidence interval of the model are shown in Figure 4.

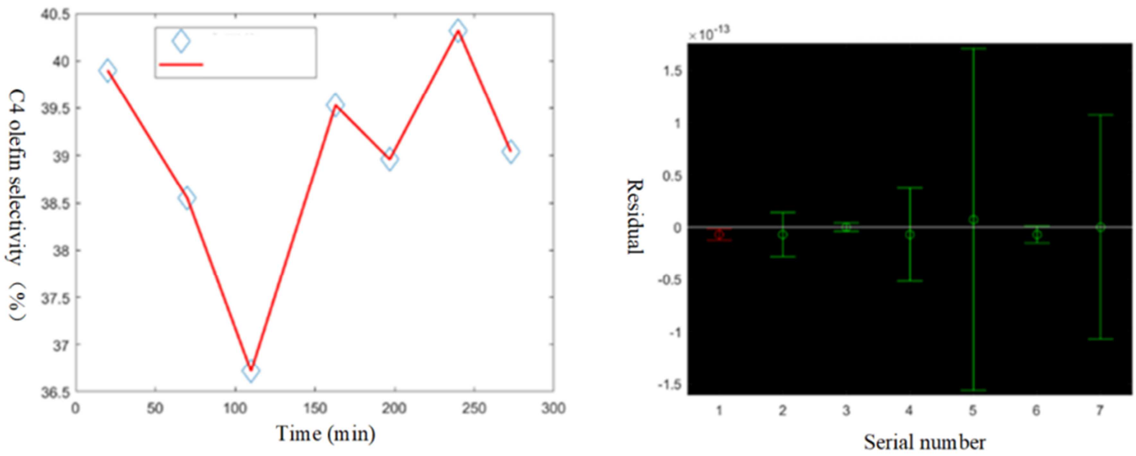

Figure 4. Comparison image between the estimated value of y_2 and the actual value and the model residual confidence interval.

Through the above modeling analysis, it can be seen that the multiple linear regression model has a good effect on the estimation of ethanol conversion and C4 olefin selectivity. The simulation accuracy of the model R^2 is greater than 0.9, and the observed residual confidence interval also contains zero, indicating that the model has a high simulation accuracy, and the model can be adopted in practical problems to solve realistic problems.

3. Results and Discussion

Based on the collected data from the process of ethanol coupling to produce C4 olefin, statistical analysis method is used to systematically analyze the data information. In Chapter 2 of this paper, we constructed regression models and exponential models for ethanol conversion, selectivity and temperature of C4 olefin, and compared and analyzed the simulation accuracy of the two models under different catalyst combinations by means of indicators such as goodness of fit, significance probability and residual variance. It was found that under the conditions of catalyst combinations A2, A3, A4, A6 and A7, the simulation results of linear regression model are better than those of exponential model, indicating that under the conditions of catalyst combination A2, A3, A4, A6 and A7, the ethanol conversion rate increases linearly with the increase of temperature. In the case of catalyst combinations A1, A5, A8, A9, A10, A11, A12, A13, A14, B1, B2, B3, B4, B5, B6 and B7, the simulation results of the exponential model are better than that of the linear regression model, indicating that the ethanol conversion rate increases exponentially with the increase of temperature under these catalyst combinations. Under the combination of catalysts A1, A3, A5, A8, A9, A11, A13, B4, B6 and B7, the modeling accuracy of the quantitative relationship between C4 olefin selectivity and temperature based on regression model is better than that of the exponential model. Under the catalyst combination A2, A4, A6, A7, A10, A12, A14, B1, B2, B3, B5, the quantitative relationship between C4 olefin selectivity and temperature based on the exponential model is better than the regression model. Through the above calculation results, we found that the linear regression model was used to model the relationship between ethanol conversion and temperature, and the change range of goodness of fit was 0.7920-1.0000, and the average goodness of fit was 0.9004; The exponential model was used to model the relationship between ethanol conversion and temperature. The variation range of goodness of fit was 0.8707-1.0000, and the average goodness of fit was 0.9672. Overall, the exponential model was more accurate to model the relationship between ethanol conversion and temperature. Linear regression model was used to model the relationship between selectivity and temperature of C4 olefin. The goodness of fit range was 0.7421-0.9888, and the average goodness of fit was 0.9175. The exponential model was used to model the relationship between selectivity and temperature of C4 olefin. The variation range of goodness of fit was 0.7787-0.9911, and the average goodness of fit was 0.9177. It can be seen that both the linear regression model and the exponential model have higher accuracy in modeling the relationship between selectivity and temperature of C4 olefin, and the exponential model has higher accuracy.

Through modeling and analysis of the relationship between ethanol conversion, C4 olefin selectivity and other products under temperature determination, we found that the use of multiple linear regression model has a good effect on the estimation of ethanol conversion and C4 olefin selectivity. The simulation accuracy of the model R^2 is greater than 0.9, and the observed residual confidence interval contains zero. It shows that the model has a high simulation accuracy. In practical application, we can estimate the ethanol conversion and C4 olefin selectivity by using the established multiple linear regression model according to the ethylene selectivity, acetaldehyde selectivity, carbon number 4-12 fatty alcohols, methyl benzaldehyde and other products.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant No. 2020YFA0608602), the National Natural Science Foundation of China (Grant No. 72243005, 72174091), Qing Lan Project of Jiangsu Province (2021), Special Science and Technology Innovation Program for Carbon Peak and Carbon

Neutralization of Jiangsu Province (Grant No. BE2022612) and the China Postdoctoral Foundation (Grant No. 2021M691312).

Conflicts of Interest

I would like to declare that they have no conflict of interests.

References

[1] FANG Yangyang, Feng Gaoyao, Le Qian et al. Calculation model for the yield of C4 olefin from chemical ethanol coupling [J]. Contemporary Chemical Research, 2022, No. 115(14): 30-32.

[2] Wu Wenjun, Zhang Zhiheng, Li Xingchen et al. Analysis of process conditions for preparation of C4 olefin by ethanol coupling based on RBF neural network model [J]. Journal of Jianghan University (Natural Science Edition), 2022, 50(03): 29-35.

[3] Li Sanjie, Hu Yulin, Wang Taihang. Preparation of C4 olefins by ethanol coupling based on grey correlation analysis [J]. Science and Technology Information, 2022, 20(14): 62-66. (in Chinese)

[4] Wang Runmo, Shen Huanyu, Chen Jing et al. Mathematical analysis of catalytic conditions in the preparation of C4 olefin by ethanol [J]. Journal of Qilu University of Technology, 2022, 36(03): 73-80. (in Chinese)

[5] Wang Liying. Data analysis of conversion rate of C4 olefin prepared by chemical industry based on MATLAB [J]. Information and Computer (Theoretical Edition), 2022, 34(11): 23-26.

[6] Zhang Yuan, Jin Yilan, Li Xiao. Synthesis of C4 olefins by ethanol coupling based on machine learning and multivariate nonlinear fitting [J]. Science and Technology Innovation, 2022(16): 177-180.

[7] JIANG Zihao, Ren Tao, Zhou Qi, et al. Optimization model of C4 olefin preparation based on Random forest and L-M algorithm [J]. Journal of Jishou University (Natural Science Edition), 2022, 43(03): 79-83. (in Chinese)

[8] Bo Xiaohan, Yan Ziqin, Wang Zhipeng et al. Preparation of C4 olefin based on multiple regression model [J]. Science and Technology Innovation, 2022(11): 49-52.

[9] LI Shaowei, Wang Yujie, Xiong Lang et al. Optimization model of ethanol coupling for preparation of C4 olefin [J]. Journal of Taizhou University, 2021, 43(06): 26-32+77. (in Chinese)

[10] Li Ming, YU Jianbo, Lu Yibo, et al. Study on preparation of C4 olefin by ethanol coupling based on regression analysis model [J]. Jiangxi Chemical Industry, 2021, 37(06): 92-94. (in Chinese)

[11] Zhang S M, Zhan W L, Hu H, et al. Research on ethanol coupling to prepare C4 olefins based on BP neural network and cluster analysis [J]. Journal of Chemistry, 2022, 2022: 1-10.

[12] Shi T. Study on Preparation of C4 olefins by ethanol coupling based on regression model [J]. International Journal of Frontiers in Engineering Technology, 2021, 3(9).

[13] Zhang Y. Application of Multiple Linear Regression Analysis in the Preparation of C4 Olefins by Ethanol Coupling [J]. Advances in Engineering Technology Research, 2022, 1(1): 168-168.

[14] Pang S, Zhu J, Zhang X. Study on Preparation of Olefins by Coupling ethanol based on partial correlation Analysis and Linear stepwise regression [J]. Highlights in Science, Engineering and Technology, 2022, 21: 383-390.

[15] Chen A. The Regression Analysis Model of C4 Olefin Yield Optimization Study [J]. Academic Journal of Materials & Chemistry, 4(5): 36-43.